BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 230 869
A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 86850416.8

(22) Date of filing: 02.12.86

(51) Int. Cl.⁴: C12N 15/00 , C12P 21/02 , C12N 1/20 , C07K 15/00

The microorganism(s) has (have) been deposited with DSM under number(s) 3594.

(30) Priority: 13.12.85 SE 8505922

(43) Date of publication of application:
05.08.87 Bulletin 87/32

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: KabiGen AB
S-112 87 Stockholm(SE)

(72) Inventor: Abrahamsén, Lars
Själagardsgatan 10
S-111 31 Stockholm(SE)
Inventor: Moks, Tomas
Kometvägen 1
S-183 33 Täby(SE)
Inventor: Nilsson, Björn
Klasrovägen 43 A
S-191 49 Sollentuna(SE)
Inventor: Uhlén, Mathias
Övre Slottsgatan 8A
S-752 35 Uppsala(SE)

(74) Representative: Burman, Tore et al
Bergling & Sundbergh AB P.O. Box 7645
S-103 94 Stockholm(SE)

(54) Construction of an IgG binding protein to facilitate downstream processing using protein engineering.

(57) A recombinant DNA fragment (Z) coding for an immunoglobulin G binding domain related to staphylococcal protein A, characterized in that the methionine codon of said fragment has been replaced by a codon of another amino acid residue enabling expression of a methioninefree protein;
a recombinant DNA sequence comprising at least two such fragments;
a recombinant DMA sequence comprising such fragment preceded by a signal sequence and followed by a nucleotide sequence coding for the amino acid sequence:
Ala Gln His Asp Glu Ala;
a recombinant DNA molecule comprising such recombinant DNA sequence and fused 3' thereof at DNA level a production gene, said molecule having the ability to express a fused protein;
a process for cleaving a fused protein expressed in a biological system by a such recombinant DNA molecule;
a plasmid vector comprising said recombinant DNA molecule; and
a bacterial cell harbouring such recombinant DNA molecule.

EP 0 230 869 A2

Construction of an IgG binding protein to facilitate downstream processing using protein engineering

The present invention relates to a recombinant DNA fragment coding for an immunoglobulin G - (hereinafter called IgG) binding domain related to staphylococcal protein A, to DNA sequences comprising such fragments and to a process for cleavage of a fused protein expressed by using such fragment or sequence. The invention also relates to plasmid vectors and bacterial cells harbouring such recombinant DNA fragments or sequences. Basically, the present invention relates to an improved system for producing and purifying foreign proteins expressed in bacteria.

Gene fusion techniques have been used in recombinant DNA technology to monitor the transient expression from a gene or to facilitate the downstream processing. By making a gene fusion to straphylococcal protein A any gene product can be purified as a fusion protein to protein A and can thus be purified in a single step using IgG affinity chromatography. We have earlier fused the protein A gene to a syntehtic gene encoding human insulin-like Growth Factor I (IGF-I). The hybrid protein expressed could be recovered in high yield from the growth medium of Staphylococcus aureus. We have also shown that a gene product consisting of divalent protein A fused to IGF-I (EE-IGF-I) was secreted from the E.coli cell by a method described in our Swedish patent application (the disclosure of which is incorporated herein by reference; filed simultaneously herewith).

These two mentioned expression systems constitute quite powerful tools for expressing and secreting foreign proteins. The use of protein A fusions, however, is dependent on processing after purification to release a biologically active peptide or protein. In an industrial process chemical cleavage methods are in preference compared to proteolytic cleavages for economical reasons. When using protein A as the carrier protein it would be of great importance if the recognition amino acid sequence would be present only in the linker to be processed to release the attached gene product so that the protein A portion could be left intact. In that way a second passage through an IgG column would bind the protein A molecule, but product released by the chemical cleavage would pass through.

For IGF-I the method suggested by us to be used is hydroxylamine cleavage in Asn-Gly dipeptide sequences. The method mostly used is otherwise CNBr cleavage specific for Met. The choice of method is dependent on if the amino acid(s) sensitive for the chemical is present in the product or not, IGF-I has an internal methionine and IGF-II has not. Protein A has, however, 3 internal methionines in the IgG binding region and 5 Asn-Gly in the IgG binding region of protein A. This makes the second passage through the column irrelevant as the protein A pieces released from the cleavage will not bind to the IgG.

The main object of this invention is to provide a solution to these problems by adapting an IgG binding domain so that no Met and optionally no Asn-Gly is present in the sequence. At the same time two non palindromic AccI sites are preferably introduced in the fragment to be able to polymerize the IgG binding domain to any number of IgG binding domains.

Accordingly, for attaining said object and other objects which will be clear from the following disclosure, the invention provides for a recombinant DNA fragment (abbreviated Z in this disclosure) coding for an immunoglobulin G binding domain related to staphylococcal protein A, such fragment being characterized in that the methionine codon of said fragment has been re placed by a codon of another amino acid residue enabling expression of a methioninefree protein. It is preferred that the codon of said another amino acid residue is that of leucine.

In a preferred embodiment of the fragment of the invention the asparagine-methionine codons have been replaced by histidineleucine codons. The codon of amino acid residue No. 1 as defined by trypsin digestion of native protein A is preferably replaced by a valine codon, so as to give at the nucleotide level the sequence GTAGAC furnishing a non-palindromic Acc I site.

In other preferred embodiment the glycine codon in the Asn-Gly constellation has been replaced by an alanine codon. The Asp-Pro codons have suitably been modified to increase the acid stability of the peptide bond of the expressed protein, such as by replacing the aspartic acid codon by a glutamic acid codon.

According to another aspect of the invention there is provided for a recombinant DNA sequence comprising at least two Z-fragments as defined above. The number of such amalgamated Z-fragments is preferably within the range 2-15, and particularly within the range 2-10.

According to still another aspect of the invention there is provided for a recombinant DNA fragment coding for any of the E D A B C domains of staphylococcal protein A, wherein the glycine codon(s) in the Asn-Gly coding constellation has been replaced by an alanine codon.

The invention also provides for a recombinant DNA sequence comprising the Z-fragment as defined above preceded by a single sequence followed by a nucleotide sequence coding for the amino acid sequence: Ala Gln His Asp Glu Ala.

The invention also covers a recombinant DNA molecule comprising the recombinant DNA sequence as described above and fused 3′ thereof at DNA level a pro duction gene. By this arrangement such molecule obtains the ability to express a fused protein in a suitable host. Such production gene may be that of a somatomedin, examples of which are: growth hormones or factors, such as hGH (human growth hormone), IGF-I, IGF-II, NGF (Nerve Growth Factor), EGF (Epithermal Growth Factor) and PDGF (Platelet Derived Growth Factor). The production gene may also be one coding for an interferon, interleuKin-2 insulin, neuropeptide, gastrointestinal peptide etc. Specifically, the production gene is that of IGF-I or IGF-II. The production gene may also code for a structural gene for an enzyme or parts thereof.

It is preferred that in such recombinant DNA molecule the N-terminal glycine codon is preceded by an asparagine codon to enable hydroxyl amino cleavage of the peptide bond to release the native protein, such as IGF-I. In another embodiment the N-terminal codon is preceded by a methionine codon to enable cyanogen bromide cleavage of the peptide bond to release native protein, such as IGF-II.

According to still another aspect of the invention there is provided a process for cleaving a fused protein expressed in a biological system by the recombinant DNA molecule as defined above. Such treatment is suitably performed by hydroxyl amine treatment when the N-terminal glycine codon is preceded by an asparagine codon. When preceded by a methionine codon the cleavage is preferably performed by cyanogen bromide treatment.

Finally, the invention covers a plasmid vector comprising the recombinant DNA molecule as described above. The invention also extends to bacterial cells harbouring the recombinant DNA-molecule defined above. The molecule can be harboured in the chromosome of the bacterial cell but may also be contained in a plasmid vector.

The host cell is for example Gram negative and is particularly constituted by an E.coli.

The invention will in the following be further illustrated by non-limiting examples with refefence to the appended drawings, wherein:

Fig. 1 shows the organization of the coding region of the protein A gen. S is the signal sequence, A-E are the IgG binding domains and X is the C-terminal region with no IgG binding activity of the encoding polypeptide;

Fig. 2 shows a comparison of the different IgG binding regions. The first line shows a suggested consensus amino acid sequence of the IgG binding regions. The boxes show the stretches of amino acids involved in the two different alpha helixes. The amino acids involved in the binding to IgG are underlined. The amino acids in the different regions are shown by — for no change compared to the consensus codon, + for no amino acid change but a silent mutation and the letter for another amino acid for amino acid changes. The amino acids are given in the one letter code;

Fig. 3 shows the nucleotide sequence of the sense strand of the synthezised Z-fragment. The cleavage region is the stretch of amino acids needed for processing of the signal sequence. The Z-region is the part of the Z-fragment coding for the IgG binding domain. The amino acid changes are underlined. The restriction enzyme recognition sequences for sites used in Examples are shown;

Fig. 4 shows the nucleotide sequence of the ZZ-IGF-I encoded by the pZZ-IGF-I plasmid vector. The regions encoding the signal peptide, the cleavage region, the two Z-regions and IGF-I are shown as well as restriction sites relevant for the construction strategy;

Fig. 5 shows the strategy described in Examples section V. The synthetic oligomers were cloned in M13 mp 18 (not shown in the figure) prior to the cloning of the Z-fragment (from Hind III to Eco RI) into pUC8. By digesting pUC8-Z with AccI the Z-region is cleaved out and by religation followed by transformation the pUC8-ZZ plasmid vector could be isolated. AMP is the gene coding for the β-lactamase gene, ori is the origin of replication for E.coli, lac Z′ is the gene coding for β-galactosidase alpha fragment and Z is the synthetic fragment;

Fig. 6 shows the cloning strategy described in Examples section VI. AMP is the gene coding for β-lactamase, S is the signal sequence, A-E are the IgG binding domains of protein A, ori is the origin of replication, Z is the synthetic fragment, IGF-I is the gene for IGF-I, F1 is the origin of replication from phage f1 and lacZ is the gene for β-galactosidase;

Fig. 7 shows the construction of pASZ1 and pASZ2 as described in Examples sections III and IV. AMP is the gene encoding the β-lactamase, F1 is the origin of replication for phage f1, S is the signal sequence, A-E are the IgG binding regions, ori is the origin of replication for E.coli and Z is the synthetic fragment;

Fig. 8 shows the strategy of the process to purify IGF-I using the method of unique Asn-Gly cleavage, as described in Examples section VII; and

Fig. 9 shows the SDS gel electrophoresis of the proteins corresponding to the different steps in the process described in Examples section VII. Lane 1 shows size markers in Kilo Daltons, lane 2 shows the hydrid protein after IgG affinity purification, lane 3 shows the result of hydroxylamine cleavage, lane 4 shows the flow through of an IgG sepharose gel of hydroxylamine digested hybridprotein and lane 5 is pure IGF-I (marker). The bands corresponding to ZZ-IGF-I, ZZ and IGF-I are shown by arrows.

By cleaving the cloned synthetic fragment using the restriction enzyme AccI followed by religation of the isolated fragment, the fragments will ligate head to tail and create a tandem repeat of fragments and in theory any number of regions can be obtained. The fragment used to illustrate the invention was constructed in the following way:

Protein A consists of two distinct regions: The IgG binding region (A-E domains) and region X having no IgG binding activity (Fig. 1). The IgG binding region consists of five homologous IgG binding domains which can be cleaved apart by trypsin treatment at protein level. The B-domain has been crystallized together with IgG and structure determined by X-ray crystallography (Deisenhofer, J., Biochemistry. 20, 2361 (1981).

The five IgG binding domains consist of approximately 58 amino acids (E is shorter and D is longer) and the amino acid sequences of the regions are shown in Fig. 2.

The fragment that was synthezied had the following features:

1) Lack of any Met;

2) Lack of Asn-Gly dipeptide sequence;

3) The fragment is potimized to be synthezied at nucleotide level to facilitate cloning to get expression;

4) The fragment can be polymerized at nucleotide level to get any number of IgG binding regions;

5) The fragment is capable of being expressed in an genetic system adapted for expression and secretion.

1) There are alltogether 3 Met in the IgG binding region of protein A. In the synthetic fragment the consensus Met (the protein A consensus amino acid sequence is shown in Fig. 2) was changed in the synthetic fragment by changing the Asn-Met to codons for a His-Leu sequence.

2) The Asn-Gly dipeptide sequence is sensitive to hydroxylamine. As this sequence is kept intact in all five IgG binding domains of protein A and as this amino acid sequence is present in the middle of an alpha helix involved in the binding to IgG (Fig. 2) there is very little chance to be successful in any amino acid change. The obvious choice to change the Asn to a Gln was analyzed by computer graphics (FRODO software, Alwin Jones, Biomedical Centre, Uppsala, Sweden) using the coordinates available from the Brookhaven Protein Data Bank (Bernstein, F.C. et al J.Mol.Biol., 112, 553-542 (1972) calculated from the X-ray crystallographic structure of protein A.

Since the Asn provides for hydrogen bonds to two other residues the change of the code for the Asn to Gln was expected to destroy both the teriary structure of protein A and the binding to IgG. Instead, however, the computer analysis surprisingly showed that the Gly in the Asn-Gly dipeptide sequence could be chariged to an Ala. This change was not obvious as glycines are among the most conserved amino acids between homologous protein sequences due to their special features. The features include flexibility around the peptide bond, the usual function to start and break alpha and beta structures and the feature as an amino acid having no net charge in the polypeptide chain. However, by simulating the Gly to Ala amino acid change in the computer we concluded that this change would not interfere with folding to protein A or binding to IgG.

3) The fragment was synthezied at DNA level in 10 separate oligomers. To facilitate cloning of the fragment one HindIII site was designed in the 5' end and an EcoRI site in the 3' end. (The nucleotide sequence of the fragment is shown in Fig. 3).

4) In order to be able to polymerize the fragment at nucleotide level a non-palindromic AccI site was introduced by changing the nucleotide sequence to GTAGAC in the code for N-terminal end of the IgG binding region. In that way an Ala codon present in all regions is changed to a Val. By this nucleotide constellation the fragment can be polymerized to any multiplicity thus altering the binding capacity of the translated product. It is not obvious if this amino acid substitution will interfere with the protein A function or not.

5) By introducing a FspI site where the signal sequense shall be attached to the fragment and by starting the coded fragment with the six amino acids unique for the E-region of protein A the fragment is capable of being attached to the protein A signal sequence by molecular cloning and the six N-terminal amino acids were introduced to ensure secretion of the fragment.

Specific embodiments of the invention will now be described in detail.

Starting materials

Bacterial hosts. Three different strains of E.coli K12 were used in the Examples: HB101 (Boyer, H.W. et al J.Mol.Biol., 41, 459-572 (1969), JM 83 (Yanisch-Perron, C. et al Gene, 33, 103-119 (1985)) and JM 103 (Messing, J. et al Methods Enzymol., 101 , 20-79 (1983)). (The strains are available at the Dept of Biochemistry and Biotechnology, Royal Insitute of Technology, Stockholm,Sweden).

Cloning vehicles: The cloning vehicles used in Examples wereM13mp18 (Yanisch-Perron, C. et al Gene, 33, 103-119 (1985)) and pUC8 (Vieva, J. et al Gene 19, 259 (1982)). The vector pHL33 is a vector derived from pEMBL19 (-) (Dente et al, Nucl.Acids Res., 11 , 1645 (1983)) and pRIT4 (Nilsson, B. et al, EMBO J. 4, 1075 (1985)) constructed in the following way:
The plasmid pRIT4 was cleaved with Taq I and after fill-in reaction with Klenow polymerase, Not-I linkers were added. The reaction mixture was cleaved with EcoRI Not I and the fragment spanning over the protein A gene was isolated. This fragment was cloned into pEMBL19 (-), where the ClaI site previously had been linked to a NotI site, by cleaving that plasmid with NotI and EcoRI. This gives a vector containing a part of the protein A gene followed by the mp19 multirestriction enzyme linker.

The vector pEX4-IGF-I is pEX (Stanley,K. et al EMBO J. 3, 1429 (1984) having the syntehtic IGF-I cloned in EcoRI to Bam HI'.

The synthetic gene encoding IGF-I has been described by Elmblad, A. et al in Third European Congress on Biotechnology III, 287-296, Verlag Chemie, Weinheim (1984). The plasmid vector pASZ2 has been deposited with the Deutsche Sammlung von Mikroorganismen (DSM), Göttingen, Federal Republic of Germany, under No. 3594 DSM and (the designated name pE'2 in the deposition document).

Buffers and Media

Coating buffer: 1.59 g Na_2CO_3, 2.93 g $NaHCO_3$ and 0.2 g NaN_3, made up to 1 litre with distilled H_2O

PBST: 8.0 g NaCl, 0.2 g KH_2PO_4, 2.9 g Na_2HPO_4 X $12H_2O$, 0.2 g KCL, 0.2 ml Tween®20 and 0.2 g NaN_3 made up to 1 litre with distilled H_2O (pH 7.4).

TSB: 30 g Tryptic Soy Broth, made up to 1 litre and autoclaved.

LB: 10 g Bacto tryptone, 5 g Yeast extract and 10 g NaCl made up to 1 litre and autoclaved.

LA: Plates containing LB and 15 g Bacto Agar per litre.

TBAB: 30 g Tryptic Blood Agar Base, made up to 1 litre and autoclaved.

ONPG-buffer 2mM o-Nitrophenyl-β-D-galactoside (ONPG, Sigma product No N-1127) in 0.1 M potassium phosphate buffer, pH 7.3, containing 15 mM 2-mercaptoethanol and 1mM $MgCl_2$.

Routine Methods

Certain procedures were carried out repeatedly in the Examples. Unless otherwise specified, they were done exactly as follows each time they were carried out.

Methods used routinely in molecular biology are not described (like the use of commercial restriction enzymes, DNA-ligations, Bal 31 exonuclease, S1 nuclease and Klenow polymerase)

Transformations: Transformation of E.coli K12 with plasmid DNA was performed exactly as described - (Morrison, D.A., Methods in Enzymology, Academic Press 68, 326-331 (1979)). The transformants were selected in a conventional manner on plates (TBAB) containing 70 mg/l ampicillin.

Isolation of plasmid DNA : Plasmid DNA was isolated as described by Birnboim, H.C. et al, Nucl.Acids Res. 7, 1513 (1979). Small scale preparations to screen a large number of transformants were made exactly as described by Kieser, T. Plasmid 12, 19-36 (1984).

Sepharose 6B chromatography: Plasmid DNA to be used for Bal31 or S1 treatment were run on a Sepharose 6B gelfiltration in a 10mM Tris, 1mM EDTA and 500 mM NaCl-buffer. In this way DNA is separated from RNA.

Elution of DNA fragments. Elution of DNA fragments from either agarose or polyacrylamide gel pieces were performed exactly as described by Maxam et al, P.N.A.S. (USA), 74, 560-564 (1977).

Ligation of DNA in Low Gel Temperature Agarose gel: Ligation directly in agarose gel was performed by running the electrophoresis in a Low Gel Temperature Agarose gel and after cutting out the band the gel piece was melted by heating to 65°C. After a 10 times dilution using Tris buffer (10 mM pH 7.4) ligation could be performed.

Detection and quantification of protein A. An ELISA test (Enzyme linked immunosorbent assay) wa used to quantify protein A. The test makes use of a special microtiter plate (Titertek, Amstelstac Netherlands) having no net charge. The wells are coated with human IgG (Kabi AB, Sweden) in a coatin buffer. Test samples are added and protein A is bound to the Fc portions of the IgG adsorbed in the wel Protein A is then assayed by an anti-protein A (from rabbit) conjugated to β-galactosidase (from Pharmaci AB, Uppsala, Sweden).

Assay: The wells of a microtiterplate are filled with 75 µl of a solution of human IgG at 16 ng/ml i Coating Buffer and the plate is incubated at room temperature for at least 1 hour. The wells are washe three times with 100 µl PBST and 50 µl of sample is added to each well. For quantitative determination fold dilutions are made. After incubation for 1 hour the wells are washed 3 times with 100 µl PBST followe by addition of 50 µl anti-protein A-β-galactosidase (the amount of protein A binding capacity added to eac well corresponds to the molar amount of IgG added to each well as detected by titration with protein A i excess). After incubation for 45 minutes, the wells were washed 3 times with 100 µl PBST followed b addition of 125 µl ONPG buffer. After incubation for 20-30 minutes 150 µl 0.1M NaOH was added to sto the reaction. The quantification is made by running a 2-fold dilution of a protein A standard solution o known concentration in parallel with the 2-fold dilutions of the test samples. The absorbance at 405 nm i measured for each well by a fotometer.

SDS-PAGE: SDS-polyacrylamide gel electrophoresis was performed exactly as described by Laemmli O.K. Nature (London), 227, 680-685 (1970) using a 10-20 % step gradient gel.

EXAMPLES

I Construction of a synthetic protein A fragment (Z)

The derived DNA sequence was analysed by a computer program and divided into 10 oligonucleotide varying in length from 41 to 45 nucleotides and with an overlap of 6 bp.

Synthesis was effected on a fully automated machine and the deprotected oligomers were purified b polyacrylamid electrophoresis (20 % polyacrylamide, 7 M Urea, 50 mM Tris-borate pH 8.3) followed b extraction into water and lyophilization.

II Ligation and cloning of Z-fragment

100 pmoles of oligonucleotides A1-A5 and B1-B5 were phosphorylated separately in 20 µl Kinase buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT) 5 units of polynucleotide kinase was added and the mixtures were incubated for 45 minutes at 37°C.

5 µg of the vector M13mp18, replicative form, was digested with the restriction enzymes EcoRI and HindIII. The large fragment from the digestion was isolated from a low temperature agarose gel.

The agarose containing digested M13 mp18 was melted at 65°C and 5 µl (0.5 µl, 0.1 p mole) was mixed with 0,5 p mole each of the phosphorylated oligomers A1-A5 and B1-B5 in 50 µl ligation buffer (66 mM Tris-HCl pH 7.6 50 mM $MgCl_2$, 50 mM DTT, 1 mM ATP) heated to 90°C and slowly cooled to room temperature during one hour. 10 units of T4 DNA ligase was added and the mixture was incubated overnight at 15°C.

E.coli JM 103 was transfected with the DNA thus obtained and grown overnight on 2 x YT plates containing x-gal and IPTG. 78 white plaques were transferred to a new 2 x YT plate and grown as colonies. Colonyhybridization with ^{32}P labelled oligonucleotide B5 as probe gave one positive colony which was picked and grown in 15 ml 2 x YT with E.coli JM 103. The cells were spun down and the phages were recovered from the supernatant. The single stranded phage DNA was extracted and purified and was used as template for sequencing reaction according to the dideoxy method. The M13mp18 containing the Z-fragment was designated M13-Z.

III Construction of the plasmid vector pASZ2

The plasmid vector PHL33 was cleaved with HindIII and EcoRI. The large fragment was isolated from a 1 % agarose gel after electrophoresis. The fragment was ligated to an isolated Z-fragment (HindIII/EcoRI). After ligation the mixture was transformed to E.coli HB101 and the isolated vector PASZ1 (Fig. 7) has the synthetic Z-fragment cloned down stream from the signal sequence. In order to place the synthetic fragment directly after the signal sequence the pASZ1 vector was cleaved with FspI. After religation the mixture was transformed into E.coli HB101 and pASZ2 could be isolated.

The vector has the synthetic Z-fragment (Fig. 3) directly after the signal sequence of protein A.

The construction was confirmed by DNA sequencing.

IV Expression of pASZ2

Expression of pASZ2 in E.coli HB101.

The strain was inoculated to 15 ml TSB and after incubation in a shake flask for 12 h the cell suspension was centrifuged.

The cells were washed once in TE (5 ml) and was further resuspended in 5 ml TSB followed by sonication for 3 x 30 s (MSE sonicator, microtip, power 6). After sonication the mixture was centrifuged 10 000 xg for 10 minutes.

The growth media and the extract from the cell sonication was assayed for protein A using the method described in Routine Methods. The results are shown in Table 1:

Plasmid	Filamentous growth	Expression level (mg/l)	% extra-cellular
pASZ2	+	3.6	72

These results show that the fragment is IgG binding.

V Construction of dimeric Z fragment

The single stranded DNA from M13-Z was annealed to sequencing primer followed by treatment with Klenow and dNTPs.

In this was double stranded DNA obtained. The reaction mixture was digested with HindIII and EcoRI and the Z fragment was isolated from a Low Gel Temperature Agarose gel electrophoresis. Separately the plasmid pUC 8 was digested with EcoRI and HindIII and the large fragment was isolated from Low Gel Temperature Agarose gel electrophoresis. The two isolated gel fragments were heated to 65°C and 2 µl of each melted fragment were ligated in 50 µl ligation buffer and 2.5 units DNA ligase at 15°C overnight. The DNA thus obtained was used to transform E .coli strain JM83 as above, and plasmid DNA from two transformants was isolated and proved to contain the Z fragment. The pUC 8-Z was digested with restriction enzyme AccI. After religation of the digestion mixture a new transformation of JM83 was made, followed by plasmid isolation from 12 transformants. Digestion with HindIII and EcoRI and analysis on agarose gel confirmed that 2 transformants carried the pUC 8 with an insert of two Z fragments (pUC 8-ZZ).

VI Construction of the expressionvector pZZ-IGF-I

The construction of pZZ-IGF-I was made in the following way (Fig. 6):

A. pUC18-ZZ was digested with FspI and EcoRI and the smallest fragment was isolated on LGT agarose.

B. The plasmid vector pHL33 was digested with FspI. The largest fragment (2273 bp) was isolated on LGT agarose.

C. The plasmid pEX4-IGF-I was digested with FspI and EcoRI. The small fragment spanning over the IGF-I gene into the AMP gene was isolated.

The three fragments from A, B and C were ligated together as described in Routing Methods and th ligation mixture was transformed into E.coli JM83.

Transformant selection was conducted using a LB agar medium containing 70 µg/ml of ampicillir Isolation of the plasmid DNA and analysis with restriction enzymes confirmed that the transformants carrie the plasmid pZZ-IGF-I.

VII Growth of E.coli JM83 containing pZZ-IGF-I and purification of IGF-I

4 l of growth medium LB supplemented with 0.2 % glucose, 0.01 M $MgSO_4$ and Ampicillin (100 µg/ml was inoculated with E.coli JM83/pZZ-IGF-I. The cells were grown in shake flasks for 20 h at 37°C in 2 flasks with 500 ml broth in each flask. The cells were spun down and kept at -20°C for 3 days. To th frozen pellet (32 g net weight) 128 ml 10 mM tris HCl pH 8.0 solution was added and the cellsuspensior was gently mixed for 2 h at +4°C. The cells were spun down by centrifugation and the supernatant wa recovered for affinity chromatography. The supernatant was recentrifuged before applied to an IgC sepharose column with 150 mM NaCl 50 mM tris HCl pH 7.5 (hereinafter TS).

The supernatant was passed through the column at a speed of 12 ml/h and the amount of IgG bindin material was analyzed before and after run through the column. The bound material was washed with TS supplemented with 0.05 % Triton X-100 and then TS and finally with 0.05 M ammonium acetate before elution with 1M acetic acid pH adjusted to 2.8 with ammonium acetate. The 2 ml fractions were assayed fo protein A content and the fractions were pooled and freeze dried. The hybrid protein was cleaved with hydroxyl amine as described by Bornstein, et al, Methods Enzymol., 47, 132 (1977). The dry material wa dissolved in 2M hydroxylamine adjusted to pH 9 with Lithium chloride and 0.2 M Tris base. The cleavage was performed at 45°C for 4 h. The cleavage reaction was stopped by lowering the pH to 7.0 with acetic acid and the material was desalted on a PD-10 column saturated with TS. The desalted fraction of 3.5 m was passed through a 1 ml IgG-sepharose column to separate IGF-I from the ZZ polypeptide. This materia was then desalted on a PD-10 column saturated with 0.2 M acetic acid, and then freeze dried. The materia was then analyzed on sodium dodecyle sulphate polyacrylamide gel electrophores. This process is showr in Fig. 8, and SDS-polyacrylamide gel is shown in Fig. 9. The nucleotide sequence of the expressed fusior protein is shown in Fig. 4.

These results shown that the ZZ fragment is functional in a process to purify IGF-I in respect to bindin IgG to facilitate purification as well as harbouring a unique Asn-Gly cleavage site to release native IGF-I anc to enable the ZZ portion to bind to the IgG gel after a second passage through the gel.

Although the invention has been exemplified in regard to purification of IGF-I it is to be noted that the invention is applicable to purification of any gene coding for a useful protein. Thus, IGF-I-purification is only one example of the usefulness of the invention and is not to be construed to limit the scope of the invention otherwise than as defined in the appended patent claims.

The Asp-Pro dipeptide constellation present in all protein A domains known to bind human IgG (Fig. 2) was designed to be changed to a Glu-Pro. This will make the normally acid labile Asp-Pro peptide bound resistant to acid treatment (like 70 % Formic Acid at 42° for 12 h) making any other introduced Asp-Pro coding linker unique. This designed amino acid change (Asp to Glu) is not shown in Examples.

Claims

1. A recombinant DNA fragment (Z) coding for an immunoglobulin G binding domain related to staphylococcal protein A, characterized in that the methionine codon of said fragment has been replaced by a codon of another amino acid residue enabling expression of a methioninefree protein.

2. The fragment of claim 1, wherein the codon of said another amino acid residue is that of leucine.

3. The fragment of claim 1 or 2, wherein the asparagine-methionine codons have been replaced by histidine-leucine codons.

4. The fragment of claim 1, 2 or 3, wherein the codon of amino acid residue number one as defined by trypsin digestion of native protein A has been replaced by a valine codon so as to give at the nucleotide level the sequence GTAGAC furnishing a non-palindromic Acc I site.

5. The fragment of claim 1, 2, 3 or 4, wherein the glycine codon has been replaced by an alanine codon.

6. The fragment of any preceding claim, wherein the Asp-Pro codons have been modified to increase the acid stability of the peptide bond of the expressed protein.

7. The fragment of claim 6, wherein the aspartic acid codon has been replaced by a glutamic acid codon.

8. A recombinant DNA sequence comprising at least two Z-fragments as defined in any preceding claim.

9. The sequence of claim 8, wherein the number of Z-fragments is within the range 2 to 15.

10. A recombinant DNA fragment coding for any of the E D A B C domains of staphylococcal protein A or functional equivalents thereof, wherein the glycine codon(s) in the Asn-Gly coding constellation has been replaced by an alanine codon.

11. A recombinant DNA sequence comprising the Z-fragment of any of claims 1-7 preceded by a signal sequence and a nucleotide sequence coding for the amino acid sequence:
Ala Gln His Asp Glu Ala.

12. A recombinant DNA sequence comprising the sequence of claim 8 or 9 preceded by a signal sequence followed by a nucleotide sequence coding for the amino acids:
Ala Gln His Asp Glu Ala.

13. A recombinant DNA molecule comprising the recombinant DNA sequence of claim 11 or 12 and fused 3' thereof at DNA level a production gene, said molecule having the ability to express a fused protein.

14. The molecule of claim 13, wherein said production gene is that of a somatomedin.

15. The molecule of claim 14, wherein said production gene is that of IGF-1.

16. The molecule of claim 15, wherein the N-terminal glycine codon is preceded by an asparagine codon to enable hydroxyl amino cleavage of the peptide bond to release native IGF-1.

17. The molecule of claim 14, wherein the production gene is that of IGF-2.

18. The molecule of claim 17, wherein the N-terminal glycine codon is preceded by a methionine codon to enable cyanogen bromide cleavage of the peptide bond to release native IGF-2.

19. A process for cleaving a fused protein expressed in a biological system by the recombinant DNA molecule of any of claims 16 to 18.

20. The process according to claim 19 in its dependence on claim 17 or 18, wherein the cleavage is performed by CNBr-treatment.

21. The process according to claim 19 in its dependence on claim 14 or 15, wherein the cleavage is performed by hydroxyl-amine treatment.

22. A plasmid vector comprising the recombinant DNA molecule of any of claims 13-18.

23. A bacterial cell harbouring the recombinant DNA molecule of any of claims 13-18.

24. A bacterial cell according to claim 23 harbour ing said molecule in its chromosome.

25. A bacterial cell according to claim 23 harbouring the plasmid vector according to claim 22.

26. A bacterial cell according to any of claims 23-25, which is Gram negative.

27. A bacterial cell according to claim 26, which is an E.coli.

28. Proteins obtained by the process of any of claims 19-21.

Fig. 1

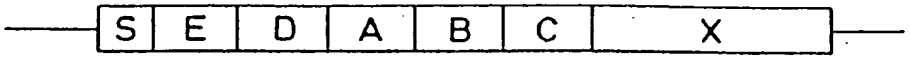


Fig. 2

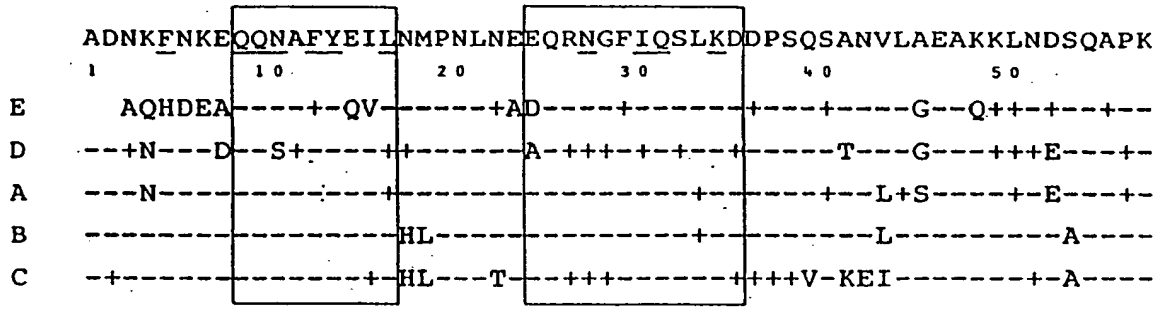


Fig. 3

Z-fragment 226 BASES

```
1HindIII  11        21 FspI   31        41AccI    51
AAGCTTCCGCGGCAAATGCTGCGCAACACGATGAAGCCGTAGACAACAAATTCAACAAA
                    AlaGlnHisAspGluAlaValAspAsnLysPheAsnLys
                      Cleavage          Z-region
                      region

60        70        80        90        110       110
GAACAACAAAACGCGTTCTATGAGATCTTACATTTACCTAACTTAAACGAAGAACAACGA
GluGlnGlnAsnAlaPheTyrGluIleLeuHisLeuProAsnLeuAsnGluGluGlnArg

120       130       140       150       160       170
AACGCCTTCATCCAAAGTTTAAAAGATGACCCAAGCCAAAGCGCTAACCTTTTAGCAGAA
AsnAlaPheIleGlnSerLeuLysAspAspProSerGlnSerAlaAsnLeuLeuAlaGlu

180       190       200       210AccI   220EcoRI
GCTAAAAAGCTAAATGATGCTCAGGCGCCGAAAGTAGACGCGAATTC
AlaLysLysLeuAsnAspAlaGlnAlaProLysValAspAlaAsn
```

ZZ-IGF-1 Fig. 4 894 BASES

```
1          11         21         31         41         51
TCGAAATAGCGTGATTTTGCGGTTTTAAGCCTTTTACTTCCTGAATAAATCTTTCAGCAA

61         71         81         91         101        111
AATATTTATTTTATAAGTTGTAAAACTTACCTTTAAATTTAATTATAAATATAGATTTTA

121        131        141        151        161        171
GTATTGCAATACATAATTCGTTATATTATGATGACTTTACAAATACATACAGGGGGTATT

181        191        201        211        221        231
AATTTGAAAAAGAAAAACATTTATTCAATTCGTAAACTAGGTGTAGGTATTGCATCTGTA
   LeuLysLysLysAsnIleTyrSerIleArgLysLeuGlyValGlyIleAlaSerVal

241        251        261        271        281        291
ACTTTAGGTACATTACTTATATCTGGTGGCGTAACACCTGCTGCAAATGCTGCGCAACAC
ThrLeuGlyThrLeuLeuIleSerGlyGlyValThrProAlaAlaAsnAlaAlaGlnHis
                                   The signal sequence      Cleavage region

301        311 AccI   321        331        341        351
GATGAAGCCGTAGACAACAAATTCAACAAAGAACAACAAAACGCGTTCTATGAGATCTTA
AspGluAlaValAspAsnLysPheAsnLysGluGlnGlnAsnAlaPheTyrGluIleLeu
                 Z-region

361        371        381        391        401        411
CATTTACCTAACTTAAACGAAGAACAACGAAACGCCTTCATCCAAAGTTTAAAAGATGAC
HisLeuProAsnLeuAsnGluGluGlnArgAsnAlaPheIleGlnSerLeuLysAspAsp

421        431        441        451        461        471
CCAAGCCAAAGCGCTAACCTTTTAGCAGAAGCTAAAAAGCTAAATGATGCTCAGGCGCCG
ProSerGlnSerAlaAsnLeuLeuAlaGluAlaLysLysLeuAsnAspAlaGlnAlaPro

481 AccI     491        501        511        521        531
AAAGTAGACAACAAATTCAACAAAGAACAACAAAACGCGTTCTATGAGATCTTACATTTA
LysValAspAsnLysPheAsnLysGluGlnGlnAsnAlaPheTYrGluIleLeuHisLeu
          Z-region

541        551        561        571        581        591
CCTAACTTAAACGAAGAACAACGAAACGCCTTCATCCAAAGTTTAAAAGATGACCCAAGC
ProAsnLeuAsnGluGluGlnArgAsnAlaPheIleGlnSerLeuLysAspAspProSer

601        611        621        631        641        651
CAAAGCGCTAACCTTTTAGCAGAAGCTAAAAAGCTAAATGATGCTCAGGCGCCGAAAGTA
GlnSerAlaAsnLeuLeuAlaGluAlaLysLysLeuAsnAspAlaGlnAlaProLysVal

661   EcoR671      681        691        701        711
GACGCGAATTCTAACGGTCCCGAAACTCTGTGCGGTGCTGAACTGGTTGACGCTCTGCAG
AspAlaAsnSerAsnGlyProGluThrLeuCysGlyAlaGluLeuValAspAlaLeuGln
                       IGF-1
721        731        741        751        761        771
TTTGTTTGCGGTGACCGTGGTTTTTATTTTAACAAACCCACTGGTTATGGTTCTTCTTCT
PheValCysGlyAspArgGlyPheTYrPheAsnLysProThrGlyTyrGlySerSerSer

781        791        801        811        821        831
CGTCGTGCTCCCCAGACTGGTATTGTTGACGAATGCTGCTTTCGTTCTTGCGACCTGCGT
ArgArgAlaProGlnThrGlyIleValAspGluCysCysPheArgSerCysAspLeuArg

841        851        861        871        881        891HindIII
CGTCTGGAAATGTATTGCGCTCCCCTGAAACCCGCTAAATCTGCTTAGAAGCTT
ArgLeuGluMetTyrCysAlaProLeuLysProAlaLysSerAla***
```

Fig. 5

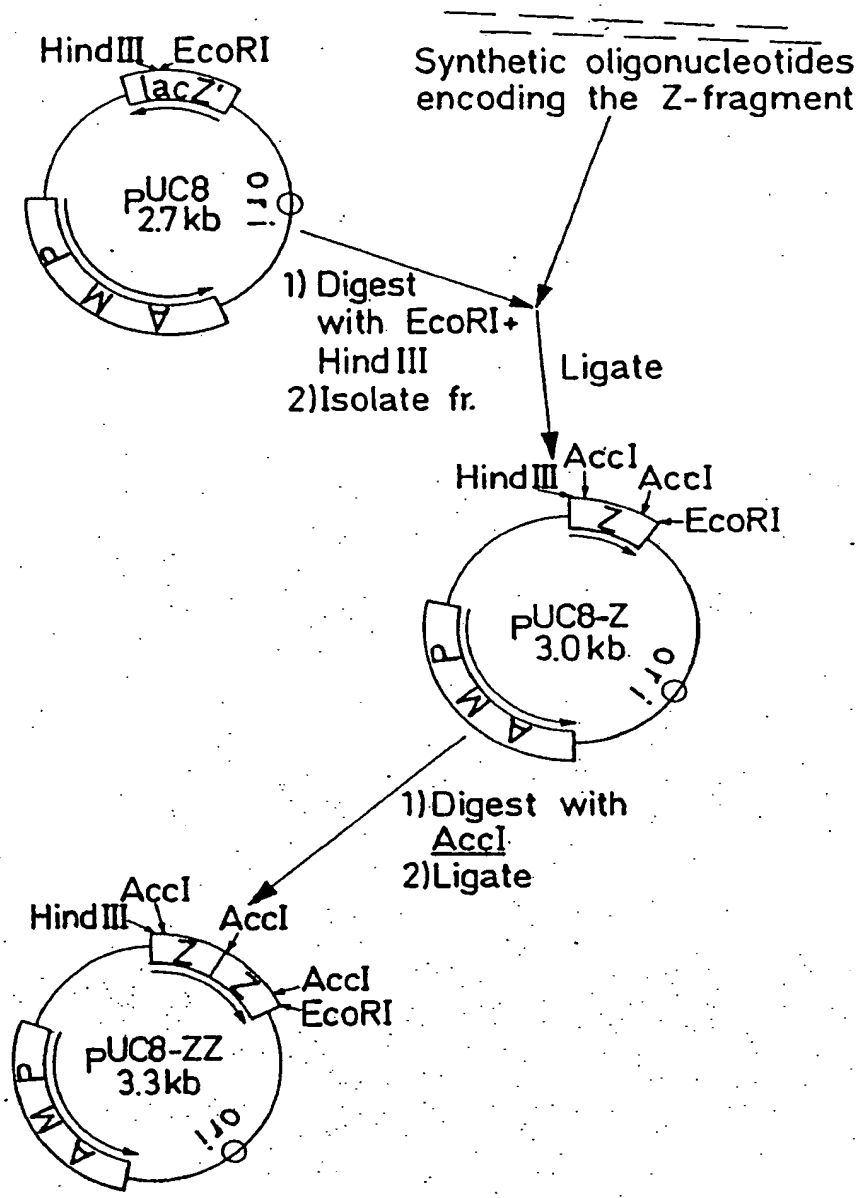

HindIII EcoRI
lacZ'
pUC8
2.7kb
ori
Amp
Synthetic oligonucleotides
encoding the Z-fragment
1) Digest
with EcoRI+
HindIII
2)Isolate fr.
Ligate
HindIII AccI AccI
Z
EcoRI
pUC8-Z
3.0kb
1) Digest with
AccI
2)Ligate
HindIII AccI AccI
AccI
EcoRI
pUC8-ZZ
3.3kb

Fig. 6

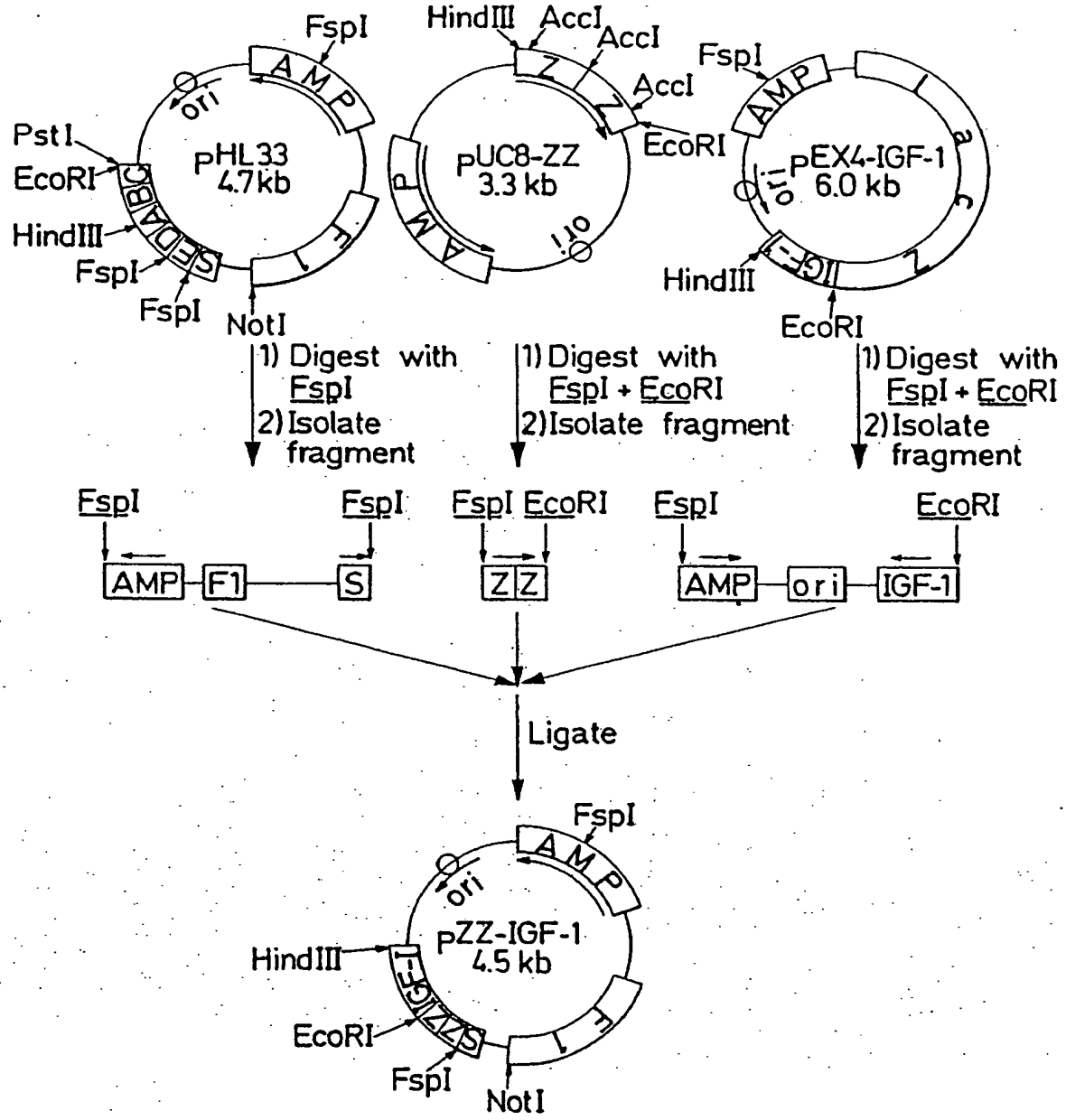

FspI
HindIII
AccI
AccI
AccI
FspI
PstI
EcoRI
HindIII
FspI
FspI
NotI
pHL33
4.7kb
pUC8-ZZ
3.3 kb
EcoRI
pEX4-IGF-1
6.0 kb
HindIII
EcoRI
1) Digest with FspI
2) Isolate fragment
1) Digest with FspI + EcoRI
2) Isolate fragment
1) Digest with FspI + EcoRI
2) Isolate fragment
FspI
FspI
FspI EcoRI
FspI
EcoRI
AMP
F1
S
Z Z
AMP
ori
IGF-1
Ligate
FspI
HindIII
EcoRI
FspI
NotI
pZZ-IGF-1
4.5 kb

Fig. 7

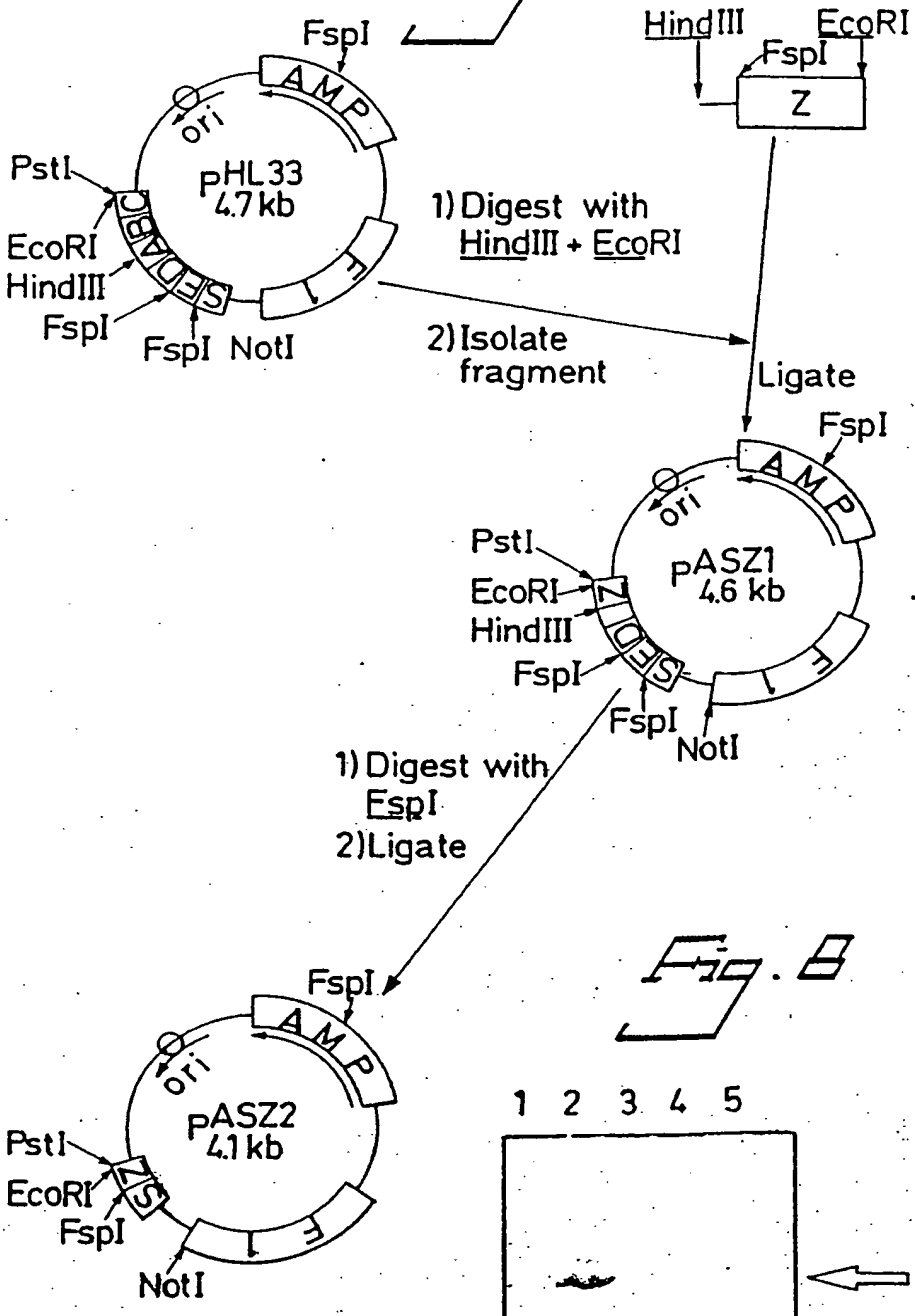

FspI
AMP
ori
pHL33
4.7 kb
PstI
EcoRI
HindIII
FspI
S E D A B C
FspI NotI
1) Digest with
HindIII + EcoRI
2) Isolate
fragment
HindIII
EcoRI
FspI
Z
Ligate
FspI
AMP
ori
pASZ1
4.6 kb
PstI
EcoRI
HindIII
FspI
S E D Z
FspI NotI
1) Digest with
FspI
2) Ligate
FspI
AMP
ori
pASZ2
4.1 kb
PstI
EcoRI
FspI
S Z
NotI

Fig. 8

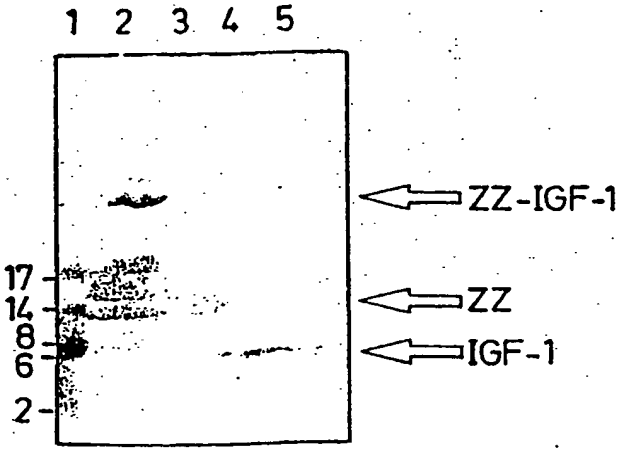
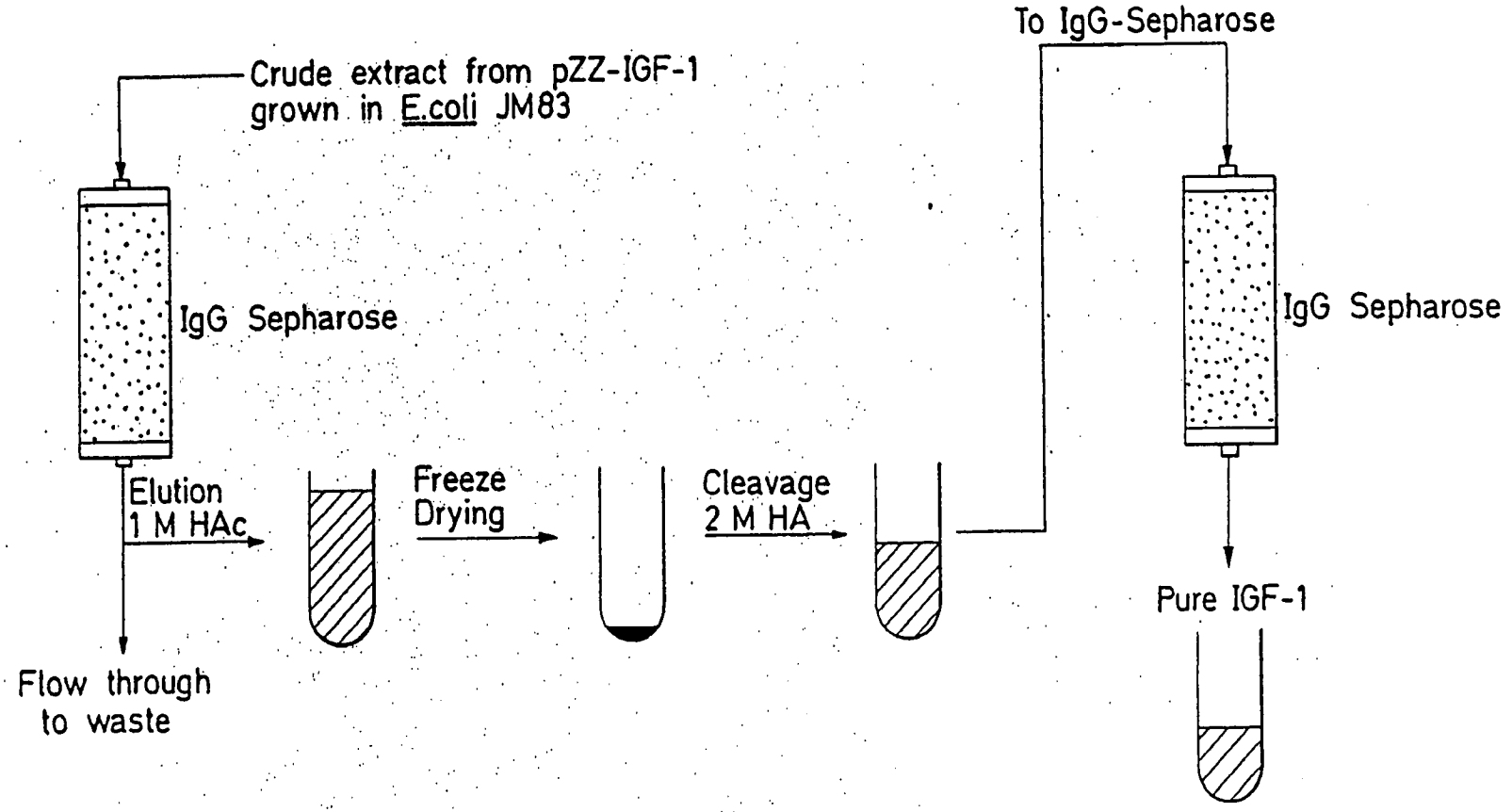

1 2 3 4 5
17
14
8
6
2
ZZ-IGF-1
ZZ
IGF-1

Fig. 9



0 230 869